



ENTERED OIPE

## RAW SEQUENCE LISTING

DATE: 03/26/2002

PATENT APPLICATION: US/10/092,880

TIME: 13:16:51

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF3\03262002\J092880.raw

3 <110> APPLICANT: Barenkamp, Stephen J.  
 5 <120> TITLE OF INVENTION: HIGH MOLECULAR WEIGHT SURFACE PROTEINS OF NON-TYPEABLE  
 6 HAEMOPHILUS  
 W--> 8 <130> FILE REFERENCE:  
 C--> 10 <140> CURRENT APPLICATION NUMBER: US/10/092,880  
 C--> 11 <141> CURRENT FILING DATE: 2002-03-08  
 13 <150> PRIOR APPLICATION NUMBER: 09/155,614  
 14 <151> PRIOR FILING DATE: 1998-09-30  
 16 <150> PRIOR APPLICATION NUMBER: 08/617,697  
 17 <151> PRIOR FILING DATE: 1996-04-01  
 19 <150> PRIOR APPLICATION NUMBER: PCT/US97/04707  
 20 <151> PRIOR FILING DATE: 1997-04-01  
 22 <160> NUMBER OF SEQ ID NOS: 11  
 24 <170> SOFTWARE: PatentIn Ver. 2.1  
 26 <210> SEQ ID NO: 1  
 27 <211> LENGTH: 5116  
 28 <212> TYPE: DNA  
 29 <213> ORGANISM: Haemophilus influenzae  
 31 <400> SEQUENCE: 1  
 32 acagcgttct cttaatacta gtacaaaccc acaataaaat atgacaaaca acaattacaa 60  
 33 cacctttttt gcagtctata tgcaaatatt ttaaaaaata gtataaatcc gccatataaa 120  
 34 atggtataat ctttcatctt tcatctttca tctttcatct ttcattcttc atctttcatc 180  
 35 tttcatcttt catctttcat ctttcatctt tcatctttca tctttcatct ttcattcttc 240  
 36 acatgccctg atgaaccgag ggaaggagg gaggggcaag aatgaagagg gagctgaacg 300  
 37 aacgcaaatg ataaagtaat ttaattgttc aactaacctt aggagaaaat atgaacaagc 360  
 38 tatatcgtct caaattcagc aaacgcctga atgcttttgt tgcgtgtgtc gaattggcac 420  
 39 ggggttgtga ccattccaca gaaaaaggca gcgaaaaacc tgctcgcatg aaagtgcgtc 480  
 40 acttagcggt aaagccactt tccgctatgt tactatcttt aggtgtaaca tctattccac 540  
 41 aatctgtttt agcaagcggc ttacaaggaa tggatgtagt acacggcaca gccactatgc 600  
 42 aagtagatgg taataaaacc attatccgca acagtgttga cgatatcatt aattggaaac 660  
 43 aatttaacat cgaccaaagt gaaatggtgc agtttttaca agaaaacaac aactccgccg 720  
 44 tattcaaccg tgttacatct aaccaaactc cccaattaaa agggatttta gattctaacg 780  
 45 gacaagtctt tttaatcaac ccaaatggta tcacaatagg taaagacgca attattaaca 840  
 46 ctaatggcct tacggcttct acgctagaca tttctaacga aaacatcaag gcgcgtaatt 900  
 47 tcaccttcga gcaaaccaaa gataaagcgc tcgctgaaat tgtgaatcac ggtttaatta 960  
 48 ctgtcggtaa agacggcagt gtaaactctta ttggtggcaa agtgaaaaac gaggggtgtg 1020  
 49 tttagcgtaaa tgggtggcagc atttctttac tcgcagggca aaaaatcacc atcagcgata 1080  
 50 taataaacc aaccattact tacagcattg ccgcgcctga aatgaagcg gtcaatctgg 1140  
 51 gcgatatttt tgccaaaggc ggtaaacatta atgtccgtgc tgccactatt cgaaccaag 1200  
 52 gtaaaccttc tgctgattct gtaagcaaag ataaaagcgg caatattggt ctttccgcca 1260  
 53 aagagggtga agcggaaatt ggcggtgtaa tttccgctca aaatcagcaa gctaaaggcg 1320  
 54 gcaagctgat gattacaggc gataaagtca cattaataac aggtgcagtt atcgacctt 1380  
 55 caggtaaaga agggggagaa acttaccttg gcggtgacga gcgcggcgaa ggtaaaaagg 1440

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/092,880

DATE: 03/26/2002

TIME: 13:16:51

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF3\03262002\J092880.raw

```

56 gcattcaatt agcaaagaaa acctctttag aaaaaggctc aaccatcaat gtatcaggca 1500
57 aagaaaaagg cggacgcgct attgtgtggg gcgatattgc gtttaattgac ggcaatatta 1560
58 acgctcaagg tagtgggtgat atcgctaaaa ccggtgggtt tgtggagacg tcgggggcatg 1620
59 atttattcat caaagacaat gcaattgttg acgccaagaa gtggttggtt gaccocggata 1680
60 atgtatctat taatgcagaa acagcaggac gcagcaatac ttcagaagac gatgaatata 1740
61 cgggatccgg gaatagtgcc agcaccccaa aacgaaacaa agaaaagaca acattaacaa 1800
62 acacaactct tgagagtata ctaaaaaaag gtacctttgt taacatcact gctaataaac 1860
63 gcattctatgt caatagctcc attaatattt ccaatggcag cttaactctt tggagtggag 1920
64 gtcggagcgg tggcggcggt gagattaaca acgatattac caccggtgat gataccagag 1980
65 gtgcaaaact aacaattttac tcaggcggct ggggtgatgt tcataaaaa atctcactcg 2040
66 gggcgcaagg taacataaac attacagcta aacaagatat cgcttttgag aaaggaagca 2100
67 accaagtcat tacagggtcaa gggactatta cctcaggcaa tcaaaaagg ttagatttta 2160
68 ataattgtct tctaaacggc actggcagcg gactgcaatt caccactaaa agaaccaata 2220
69 aatacgctat cacaaataaa tttgaaggga ctttaaatat ttcagggaaa gtgaacatct 2280
70 caatggtttt acctaaaaat gaaagtggat atgataaatt caaaggacgc acttactgga 2340
71 atttaacctc cttaaatggt tccgagagtg gcgagtttaa cctcactatt gactccagag 2400
72 gaagcgatag tgcaggcaca cttaccacgc cttataatth aaacgggata tcattcaaca 2460
73 aagacactac ctttaattgt gaacgaaatg caagagtcaa ctttgacatc aaggcaccaa 2520
74 tagggataaa taagtattct agtttgaatt acgcatcatt taatggaaac atttcagttt 2580
75 cgggaggggg gagtggtgat ttcacacttc tcgcctcatc ctctaacgta caaacccccg 2640
76 gtgtagttat aaattctaaa tacttttaag tttaacagg gtcaagttta agatttaaaa 2700
77 cttcaggctc acaaaaaact ggcttctcaa tagaagaaag ttttaacttt aatgccaccg 2760
78 gaggaacat aacacttttg caagttgaag gcaccgatgg aatgatttgt aaaggcattg 2820
79 tagccaaaaa aacataaacc tttgaaggag gtaacatcac ctttggctcc aggaaagccg 2880
80 taacagaaat cgaaggcaat gttactatca ataacaacgc taacgtcact cttatcggtt 2940
81 cggattttga caaccatcaa aaacctttta ctttaaaaa agatgtcatc attaatagcg 3000
82 gcaaccttac cgctggaggc aatattgtca atatagccgg aaatcttacc gttgaaagta 3060
83 acgctaattt caaagctatc acaattttca cttttaatgt agggcgcttg tttgacaaca 3120
84 aaggcaattc aaatatttcc attgccaaag gaggggctcg ctttaaaagc attgataatt 3180
85 ccaagaattt aagcatcacc accaactcca gctccactta ccgcactatt ataagcggca 3240
86 atataaccaa taaaaacggt gatttaataa ttacgaacga aggtagtgat actgaaatgc 3300
87 aaattggcgg cgatgtctcg caaaaagaag gtaatctcac gatttcttct gacaaaatca 3360
88 atattaccaa acagataaca atcaaggcag gtgttgatgg ggagaattcc gattcagacg 3420
89 cgacaaacaa tgccaatcta accattaaaa ccaaagaatt gaaattaacg caagacctaa 3480
90 atatttcagg tttcaataaa gcagagatta cagctaaaga tggtagtgat ttaactattg 3540
91 gtaacaccaa tagtgctgat ggtactaatg ccaaaaaagt aacctttaac caggttaaag 3600
92 attcaaaaat ctctgctgac ggtcacaagg tgacactaca cagcaaagtg gaaacatccg 3660
93 gtagtaataa caacactgaa gatagcagtg acaataatgc cggtttaact atcgatgcaa 3720
94 aaaatgtaac agtaaacac aatattactt ctcaaaaagc agtgagcatc tctgcgacaa 3780
95 gtggagaaat taccactaaa acaggtacaa ccattaacgc aaccactggt aacgtggaga 3840
96 taaccgctca aacaggtagt atcctagggt gaattgagtc cagctctggc tctgtaacac 3900
97 ttactgcaac cgagggcgct cttgctgtaa gcaatatttc gggcaacacc gttactgtta 3960
98 ctgcaaatag cgggtgcatta accacttttg caggctctac aattaaagga accgagagtg 4020
99 taaccacttc aagtcaatca ggcgatatcg gcggtacgat ttctggtggc acagtagagg 4080
100 ttaaagcaac cgaagttta accactcaat ccaattcaaa aattaaagca acaacaggcg 4140
101 aggctaactg acaagtgca acaggtaaca ttggtgttac gatttcgggt aatacggtaa 4200
102 atgttacggc aaacgctggc gatttaacag ttgggaatgg cgcagaaatt aatgcgacag 4260
103 aaggagctgc aaccttaact acatcatcgg gcaaatcaac taccgaagct agttcacaca 4320
104 ttacttcagc caagggtcag gtaaatcttt cagctcagga tggtagcggt gcaggaaagta 4380

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/092,880

DATE: 03/26/2002

TIME: 13:16:51

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF3\03262002\J092880.raw

```

105 ttaatgccgc caatgtgaca ctaaatacta caggcacttt aactaccgtg aaggggttcaa 4440
106 acattaatgc aaccagcggg accttggtta ttaacgcaaa agacgctgag ctaaatggcg 4500
107 cagcattggg taaccacaca gtggtaaatg caaccaacgc aaatggctcc ggcagcgtaa 4560
108 tcgcgacaac ctcaagcaga gtgaacatca ctggggattt aatcacaata aatggattaa 4620
109 atatcatttc aaaaaacggg ataaacaccg tactgttaaa aggcgttaaa attgatgtga 4680
110 aatacattca accgggtata gcaagcgtag atgaagtaat tgaagcgaaa cgcacccctg 4740
111 agaaggtaaa agatttatct gatgaagaaa gagaagcggt agctaaactt ggagtaagtg 4800
112 ctgtacgttt tattgagcca aataatacaa ttacagtcga tacacaaaat gaatttgcaa 4860
113 ccagaccatt aagtogaata gtgattttctg aaggcagggc gtgtttctca aacagtgatg 4920
114 gcgcgacggg gtgcgttaat atcgctgata acgggcggta gcggtcagta attgacaagg 4980
115 tagatttcat cctgcaatga agtcatttta tttcgtatt atttactgtg tgggttaaag 5040
116 ttcagtacgg gctttaccca tcttgtaaaa aattacggag aatacaataa agtattttta 5100
117 acaggttatt attatg 5116
120 <210> SEQ ID NO: 2
121 <211> LENGTH: 1536
122 <212> TYPE: PRT
123 <213> ORGANISM: Haemophilus influenzae
125 <400> SEQUENCE: 2
126 Met Asn Lys Ile Tyr Arg Leu Lys Phe Ser Lys Arg Leu Asn Ala Leu
127 1 5 10 15
129 Val Ala Val Ser Glu Leu Ala Arg Gly Cys Asp His Ser Thr Glu Lys
130 20 25 30
132 Gly Ser Glu Lys Pro Ala Arg Met Lys Val Arg His Leu Ala Leu Lys
133 35 40 45
135 Pro Leu Ser Ala Met Leu Leu Ser Leu Gly Val Thr Ser Ile Pro Gln
136 50 55 60
138 Ser Val Leu Ala Ser Gly Leu Gln Gly Met Asp Val Val His Gly Thr
139 65 70 75 80
141 Ala Thr Met Gln Val Asp Gly Asn Lys Thr Ile Ile Arg Asn Ser Val
142 85 90 95
144 Asp Ala Ile Ile Asn Trp Lys Gln Phe Asn Ile Asp Gln Asn Glu Met
145 100 105 110
147 Val Gln Phe Leu Gln Glu Asn Asn Asn Ser Ala Val Phe Asn Arg Val
148 115 120 125
150 Thr Ser Asn Gln Ile Ser Gln Leu Lys Gly Ile Leu Asp Ser Asn Gly
151 130 135 140
153 Gln Val Phe Leu Ile Asn Pro Asn Gly Ile Thr Ile Gly Lys Asp Ala
154 145 150 155 160
156 Ile Ile Asn Thr Asn Gly Phe Thr Ala Ser Thr Leu Asp Ile Ser Asn
157 165 170 175
159 Glu Asn Ile Lys Ala Arg Asn Phe Thr Phe Glu Gln Thr Lys Asp Lys
160 180 185 190
162 Ala Leu Ala Glu Ile Val Asn His Gly Leu Ile Thr Val Gly Lys Asp
163 195 200 205
165 Gly Ser Val Asn Leu Ile Gly Gly Lys Val Lys Asn Glu Gly Val Ile
166 210 215 220
168 Ser Val Asn Gly Gly Ser Ile Ser Leu Leu Ala Gly Gln Lys Ile Thr
169 225 230 235 240
171 Ile Ser Asp Ile Ile Asn Pro Thr Ile Thr Tyr Ser Ile Ala Ala Pro

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/092,880

DATE: 03/26/2002

TIME: 13:16:51

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF3\03262002\J092880.raw

172		245		250		255
174	Glu Asn Glu	Ala Val Asn	Leu Gly Asp	Ile Phe Ala	Lys Gly Gly	Asn
175		260		265		270
177	Ile Asn Val	Arg Ala Ala	Thr Ile Arg	Asn Gln Gly	Lys Leu Ser	Ala
178		275		280		285
180	Asp Ser Val	Ser Lys Asp	Lys Ser Gly	Asn Ile Val	Leu Ser Ala	Lys
181		290		295		300
183	Glu Gly Glu	Ala Glu Ile	Gly Gly Val	Ile Ser Ala	Gln Asn Gln	Gln
184	305		310		315	320
186	Ala Lys Gly	Gly Lys Leu	Met Ile Thr	Gly Asp Lys	Val Thr Leu	Lys
187		325		330		335
189	Thr Gly Ala	Val Ile Asp	Leu Ser Gly	Lys Glu Gly	Gly Glu Thr	Tyr
190		340		345		350
192	Leu Gly Gly	Asp Glu Arg	Gly Glu Gly	Lys Asn Gly	Ile Gln Leu	Ala
193		355		360		365
195	Lys Lys Thr	Ser Leu Glu	Lys Gly Ser	Thr Ile Asn	Val Ser Gly	Lys
196		370		375		380
198	Glu Lys Gly	Gly Arg Ala	Ile Val Trp	Gly Asp Ile	Ala Leu Ile	Asp
199	385		390		395	400
201	Gly Asn Ile	Asn Ala Gln	Gly Ser Gly	Asp Ile Ala	Lys Thr Gly	Gly
202		405		410		415
204	Phe Val Glu	Thr Ser Gly	His Asp Leu	Phe Ile Lys	Asp Asn Ala	Ile
205		420		425		430
207	Val Asp Ala	Lys Glu Trp	Leu Leu Asp	Phe Asp Asn	Val Ser Ile	Asn
208		435		440		445
210	Ala Glu Thr	Ala Gly Arg	Ser Asn Thr	Ser Glu Asp	Asp Glu Tyr	Thr
211		450		455		460
213	Gly Ser Gly	Asn Ser Ala	Ser Thr Pro	Lys Arg Asn	Lys Glu Lys	Thr
214	465		470		475	480
216	Thr Leu Thr	Asn Thr Thr	Leu Glu Ser	Ile Leu Lys	Lys Gly Thr	Phe
217		485		490		495
219	Val Asn Ile	Thr Ala Asn	Gln Arg Ile	Tyr Val Asn	Ser Ser Ile	Asn
220		500		505		510
222	Leu Ser Asn	Gly Ser Leu	Thr Leu Trp	Ser Glu Gly	Arg Ser Gly	Gly
223		515		520		525
225	Gly Val Glu	Ile Asn Asn	Asp Ile Thr	Thr Gly Asp	Asp Thr Arg	Gly
226		530		535		540
228	Ala Asn Leu	Thr Ile Tyr	Ser Gly Gly	Trp Val Asp	Val His Lys	Asn
229	545		550		555	560
231	Ile Ser Leu	Gly Ala Gln	Gly Asn Ile	Asn Ile Thr	Ala Lys Gln	Asp
232		565		570		575
234	Ile Ala Phe	Glu Lys Gly	Ser Asn Gln	Val Ile Thr	Gly Gln Gly	Thr
235		580		585		590
237	Ile Thr Ser	Gly Asn Gln	Lys Gly Phe	Arg Phe Asn	Asn Val Ser	Leu
238		595		600		605
240	Asn Gly Thr	Gly Ser Gly	Leu Gln Phe	Thr Thr Lys	Arg Thr Asn	Lys
241		610		615		620
243	Tyr Ala Ile	Thr Asn Lys	Phe Glu Gly	Thr Leu Asn	Ile Ser Gly	Lys
244	625		630		635	640

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/092,880

DATE: 03/26/2002

TIME: 13:16:51

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF3\03262002\J092880.raw

```

246 Val Asn Ile Ser Met Val Leu Pro Lys Asn Glu Ser Gly Tyr Asp Lys
247           645           650           655
249 Phe Lys Gly Arg Thr Tyr Trp Asn Leu Thr Ser Leu Asn Val Ser Glu
250           660           665           670
252 Ser Gly Glu Phe Asn Leu Thr Ile Asp Ser Arg Gly Ser Asp Ser Ala
253           675           680           685
255 Gly Thr Leu Thr Gln Pro Tyr Asn Leu Asn Gly Ile Ser Phe Asn Lys
256           690           695           700
258 Asp Thr Thr Phe Asn Val Glu Arg Asn Ala Arg Val Asn Phe Asp Ile
259 705           710           715           720
261 Lys Ala Pro Ile Gly Ile Asn Lys Tyr Ser Ser Leu Asn Tyr Ala Ser
262           725           730           735
264 Phe Asn Gly Asn Ile Ser Val Ser Gly Gly Gly Ser Val Asp Phe Thr
265           740           745           750
267 Leu Leu Ala Ser Ser Ser Asn Val Gln Thr Pro Gly Val Val Ile Asn
268           755           760           765
270 Ser Lys Tyr Phe Asn Val Ser Thr Gly Ser Ser Leu Arg Phe Lys Thr
271           770           775           780
273 Ser Gly Ser Thr Lys Thr Gly Phe Ser Ile Glu Lys Asp Leu Thr Leu
274 785           790           795           800
276 Asn Ala Thr Gly Gly Asn Ile Thr Leu Leu Gln Val Glu Gly Thr Asp
277           805           810           815
279 Gly Met Ile Gly Lys Gly Ile Val Ala Lys Lys Asn Ile Thr Phe Glu
280           820           825           830
282 Gly Gly Asn Ile Thr Phe Gly Ser Arg Lys Ala Val Thr Glu Ile Glu
283           835           840           845
285 Gly Asn Val Thr Ile Asn Asn Ala Asn Val Thr Leu Ile Gly Ser
286           850           855           860
288 Asp Phe Asp Asn His Gln Lys Pro Leu Thr Ile Lys Lys Asp Val Ile
289 865           870           875           880
291 Ile Asn Ser Gly Asn Leu Thr Ala Gly Gly Asn Ile Val Asn Ile Ala
292           885           890           895
294 Gly Asn Leu Thr Val Glu Ser Asn Ala Asn Phe Lys Ala Ile Thr Asn
295           900           905           910
297 Phe Thr Phe Asn Val Gly Gly Leu Phe Asp Asn Lys Gly Asn Ser Asn
298           915           920           925
300 Ile Ser Ile Ala Lys Gly Gly Ala Arg Phe Lys Asp Ile Asp Asn Ser
301           930           935           940
303 Lys Asn Leu Ser Ile Thr Thr Asn Ser Ser Ser Thr Tyr Arg Thr Ile
304 945           950           955           960
306 Ile Ser Gly Asn Ile Thr Asn Lys Asn Gly Asp Leu Asn Ile Thr Asn
307           965           970           975
309 Glu Gly Ser Asp Thr Glu Met Gln Ile Gly Gly Asp Val Ser Gln Lys
310           980           985           990
312 Glu Gly Asn Leu Thr Ile Ser Ser Asp Lys Ile Asn Ile Thr Lys Gln
313           995           1000           1005
315 Ile Thr Ile Lys Ala Gly Val Asp Gly Glu Asn Ser Asp Ser Asp Ala
316 1010           1015           1020
318 Thr Asn Asn Ala Asn Leu Thr Ile Lys Thr Lys Glu Leu Lys Leu Thr

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/092,880

DATE: 03/26/2002

TIME: 13:16:52

Input Set : A:\Sequence Listing.txt

Output Set: N:\CRF3\03262002\J092880.raw

L:8 M:201 W: Mandatory field data missing, FILE REFERENCE

L:10 M:270 C: Current Application Number differs, Replaced Application Number

L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date